UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/738,987	12/19/2000	Nelly Tarbouriech	D/A0471	2709

7590 04/07/2004

OLIFF & BERRIDGE PLC
P O BOX 19928
ALEXANDRIA, VA 22320

EXAMINER
HARPER, V PAUL

ART UNIT	PAPER NUMBER
2654	

DATE MAILED: 04/07/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/738,987	TARBOURIECH ET AL.	
	Examiner	Art Unit	
	V. Paul Harper	2654	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-19* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-19* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *14 February 2002* is/are: a)☐ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *11/19/01*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Information Disclosure Statement

1. The Examiner has considered the references listed in the Information Disclosure Statement dated 10/19/2001. A copy of the Information Disclosure Statement is attached to this office action.

Drawings

2. Figure 6 should be designated by a legend such as --Prior Art-- because only that which is old is illustrated. See MPEP § 608.02(g). A proposed drawing correction or corrected drawings are required in reply to the Office action to avoid abandonment of the application. The objection to the drawings will not be held in abeyance.

3. The drawings were received on 2/14/2002. Figures 1-5 are acceptable; Figure 6 should be amended as indicated above.

Claim Rejections - 35 USC § 102

> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

4. Claims 1-4 and 8-19 rejected under 35 U.S.C. 102(b) as being anticipated by Ratnaparkhi ("A Maximum Entropy Model for Part-Of-Speech Tagging," in Proceedings of the conference on empirical methods in Natural Language Processing, pages 133-142, 1996), hereinafter referred to as Ratnaparkhi.

Regarding claim 1, Ratnaparkhi teaches a method for assigning part-of-speech tags to text (p. 133, "Introduction"); Ratnaparkhi's method includes the following:

- obtaining an identifier for the phrase, the identifier being associated with context information (abstract; p. 133, col. 2, ¶2; tags for each word—identifiers--are associated with context, p. 136, col. 1, ¶'s3 and 4; p. 141, col. 1, ¶1, probability calculation can be performed over any structure ..., such as noun phrases);
- supplementing the phrase with the context information (p. 133, col. 2, ¶2, *H* is a set of possible word contexts where the context information is contained in separate—supplemental—corpus $Introduction); and
- assigning the at least one POS-tag to the phrase based on the supplemented phrase (p. 136, § Search Algorithm, determine the highest probability sequence, i.e., each word gets assigned a tag).

Regarding claim 2, Ratnaparkhi teaches everything claimed, as applied above (see claim 1); in addition, Ratnaparkhi teaches "the context information comprises at least one of pre-context and post-context information" (p. 133, "Introduction" approach combines diverse forms of contextual information; Tables 1-4; e.g. p. 138, col. 1, ¶3, surrounding words, in particular see "about" in Table 3).

Regarding claim 3, Ratnaparkhi teaches everything claimed, as applied above (see claim 1); in addition, Ratnaparkhi teaches "the identifier is a main grammatical category

of the phrase on the whole" (Abstract, predicts POS tag; Introduction; p. 141, col. 1, ¶1 "any structure that is predicted over the POS tags ..., such as noun phrases").

Regarding claim 4, Ratnaparkhi teaches everything claimed, as applied above (see claim 1); in addition, Ratnaparkhi teaches "the identifier is defined by a structural property of the phrase in a plurality of phrases or textual information" (p. 133, col. 2, ¶2, the probability model is defined over possible word and tag contexts; p. 141, col. 1, ¶1 "any structure that is predicted over the POS tags ..., such as noun phrases").

Regarding claim 8, Ratnaparkhi teaches everything claimed, as applied above (see claim 1); in addition, Ratnaparkhi teaches the following steps:

- preselecting probable main grammatical categories for the phrase according to the identifier (p. 136, col. 1, ¶4, the search procedure generates tags for known words);
- selecting a most probable category from the preselected main grammatical categories, the most probable category being associated to the context information to be supplemented to the phrase (p. 136, §Testing the Model, chose the tag sequence for the sentence having the highest probability, i.e., select tag(s) for words, and phrases p. 141, col. 1, ¶1).

Regarding claim 9, Ratnaparkhi teaches everything claimed, as applied above (see claim 8); in addition, Ratnaparkhi teaches "the step of selecting the most probable category comprises obtaining an external selection of the most probable category" (§p.

136, §"Testing the Model" and §"Specialized Features and Consistency," specialized features are constructed for "difficult" words—external selection).

Regarding claim 10, Ratnaparkhi teaches everything claimed, as applied above (see claim 8); in addition Ratnaparkhi teaches the following:

- the at least one POS-tag assigned to the phrase is selected from potential POS-tags assignable to the phrase without context information (p. 136, generates all tags in a set for unknown words, i.e., words not in the dictionary and not in the training set and thus having no context);
- and selecting the most probable category is supported by evaluating the potential POS-tags (p. 136, Search Algorithm).

Regarding claim 11, Ratnaparkhi teaches everything claimed, as applied above (see claim 1); in addition, Ratnaparkhi teaches "the context information includes at least one POS-tag" (p. 133, §"The Probability Model" includes the set of possible word and tag contexts; p. 136, ¶4).

Regarding claim 12, Ratnaparkhi teaches everything claimed, as applied above (see claim 1); in addition, Ratnaparkhi teaches "the context information includes textual information" (§"The Probability Model" includes the set of possible *word* and tag contexts; Tables 1-4; derived from corpus, §Introduction last ¶).

Regarding claim 13, Ratnaparkhi teaches everything claimed, as applied above (see claim 1); furthermore, a statistical model as described by Ratnaparkhi will inherently be implemented as processor-executable instructions on a system with a processor connected to primary storage that assigns part-of-speech tags to previously unseen text (i.e., inputted text, §Introduction, ¶1). In addition, Ratnaparkhi teaches:

- a context supplementer for supplementing the context information to the phrase (p. 133, col. 2, ¶2, *H* is a set of possible word contexts; derived from corpus, §Introduction last ¶);
- a POS-tagger for assigning the at least one tag to the phrase (p. 136, § Search Algorithm, determine the highest probability sequence, i.e., assign most likely tags).;
- a context storage comprising a plurality of context information items (§Introduction, "Penn Treebank Wall St. Journal corpus"; p. 136, §"Search Algorithm," Tag Dictionary);
- and an identifier storage comprising a plurality of identifiers (§"The Probability Model," a set of allowable tags);
- each of which being associated with at least one context information item of the plurality of context information items (§"The Probability Model," tag contexts, or 'histories,' and a set of allowable tags);
- identifier input means for obtaining the identifier for the phrase, the identifier being associated to the context information according to the plurality of identifiers and its association to the plurality of context information items (p. 136, §"Search Algorithm" for each word known tags are given).

Regarding claim 14, Ratnaparkhi teaches a method for assigning part-of-speech tags to text (p. 133, "Introduction"). Ratnaparkhi's method includes the following:

- acquiring the phrase (§Introduction, ¶1, previously unseen text). ;
- acquiring an identifier for the phrase, which is associated with artificial information, the artificial information supporting grammatical disambiguation of the phrase (abstract; p. 133, col. 2, ¶2; tags for each word p. 136, col. 1, ¶4; p. 141, col. 1, ¶1, probability calculation can be performed over any structure ..., such as noun phrases; the artificial information is contained in the corpus, §Introduction);
- supplementing the phrase with the artificial information (§Introduction, the use of context, i.e. associated the word with words from the corpus, §"Search Algorithm");
- and grammatically disambiguating the phrase based on the supplemented phrase (§Introduction, accurate assignment of POS tags, §"Search Algorithm," i.e, find the highest probability sequence).

Regarding claim 15, Ratnaparkhi teaches an algorithmic technique for assigning part-of-speech tags to text (p. 133, "Introduction"). Ratnaparkhi's algorithmic technique is inherently executable on a processor and includes the following:

- identifier input means for obtaining an identifier for the phrase, the identifier being associated to context information (§Introduction, §"Search Algorithm" words are associated with tags);

- a context storage comprising a plurality of context information items (§Introduction, Penn Treebank Wall St. Journal corpus; §"Search Algorithm" tag dictionary);
- an identifier storage, connected to the context storage, comprising a plurality of identifiers (§"The Probability Model" possible word and tag contexts, or 'histories', and a set of allowable tags; §"Search Algorithm" identifiers correspond to a sentence or test sequence, i.e., the *w's*);
- each of which being associated with at least one context information item of the plurality of context information items (§"The Probability Model" possible word and tag contexts, or 'histories', and a set of allowable tags);
- a context supplementer, connected to the identifier input and the context storage, for supplementing the phrase with the associated context information of the obtained identifier (§"Testing the Model" Wall St. Journal corpus and tag dictionary);
- and a POS-tagger, connected to the context supplementer, for identifying the part of speech of each part of the phrase based on the supplemented phrase (§"Search Algorithm," generate the tags).

Regarding claim 16, Ratnaparkhi teaches everything claimed, as applied above (see claim 15); in addition, Ratnaparkhi teaches the use of "phrase input means, connected to the context supplementer, for getting the phrase" (§Introduction, assign POS tags to unseen text—phrase with inherent input means; §"The Probability Model" containing the set of possible word and tag contexts; and §"Search Algorithm").

Regarding claim 17, Ratnaparkhi teaches everything claimed, as applied above (see claim 15). Ratnaparkhi further teaches:

- a category storage, connected to the identifier storage, for storing a plurality of categories, the identifier being associated with at least one of the plurality of categories (§Introduction, §"The Probability Model" text input, words and tags, inherently stored during processing);
- a category evaluator, connected to the category storage, the identifier storage and the context supplementer, for evaluating the category of the phrase based on the identifier (§"The Search Algorithm," determines the highest probability sequence).

Regarding claim 18, Ratnaparkhi teaches everything claimed, as applied above (see claim 15); in addition, Ratnaparkhi teaches an "category input means connected to the context supplementer for obtaining a category associated to one of the context information items, which is supplemented to the phrase" (§"The Probability Model" word and tag contexts and allowable tags).

Regarding claim 19, Ratnaparkhi teaches everything claimed, as applied above (see claim 15); in addition, Ratnaparkhi teaches "comprising output means connected to the POS-tagger for outputting the at least one POS-tag" (Abstract, §Introduction, prediction of POS tag).

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

5. Claims 5-7 are rejected under 35 U.S.C. 103(a) as being unpatentable over Ratnaparkhi in view of Ratnaparkhi ("Learning to Parse Natural Language with Maximum Entropy Models," Machine Learning, p. 151-175, Feb 1999), hereinafter referred to as Ratnaparkhi-99.

Regarding claim 5, Ratnaparkhi teaches everything claimed, as applied above (see claim 1), but Ratnaparkhi does not specifically teach "deriving a formal structure for the phrase, that covers variations of the phrase" However, the examiner contends that this concept was well known in the art, as taught by Ratnaparkhi-99.

In the same field of endeavor, Ratnaparkhi-99 teaches the parsing of natural language with maximum entropy models that includes the use of a part-of-speech tagger that does not need to commit to a single tag sequence (p. 5, §3.1.1, Fig. 2-4).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Ratnaparkhi by specifically implementing a part-of-speech tagger, as taught by Ratnaparkhi-99, since this technique can be used to implement superior parsers (Introduction).

Regarding claim 6, Ratnaparkhi teaches everything claimed, as applied above (see claim 1), but Ratnaparkhi does not specifically teach "extracting a headword out of the phrase based on the phrase with the at least one assigned POS-tag." However, the examiner contends that this concept was well known in the art, as taught by Ratnaparkhi-99.

In the same field of endeavor, Ratnaparkhi-99 teaches the parsing of natural language with maximum entropy models that includes the use of a part-of-speech tagger to aid in the creation of a parse tree with a headword (Fig. 2-5, §3.2.2. end of ¶3).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Ratnaparkhi by specifically implementing a part-of-speech tagger, as taught by Ratnaparkhi-99, since this technique can be used to implement superior parsers (Introduction).

Regarding claim 7, Ratnaparkhi teaches everything claimed, as applied above (see claim 1), but Ratnaparkhi does not specifically teach "selecting a main grammatical category of the phrase on the whole according to the identifier, the main grammatical category being associated to the context information to be supplemented to the phrase." However, the examiner contends that this concept was well known in the art, as taught by Ratnaparkhi-99.

In the same field of endeavor, Ratnaparkhi-99 teaches the parsing of natural language with maximum entropy models that includes the use of a part-of-speech tagger to aid in the creation of a parse tree with constituents such as noun phrases (Fig. 2-5, §3.2.2. end of ¶3).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Ratnaparkhi by specifically implementing a part-of-speech tagger, as taught by Ratnaparkhi-99, since this technique can be used to implement superior parsers (Introduction).

Citation of Pertinent Art

6. The following prior art made of record but not relied upon is considered pertinent to the applicant's disclosure:

- Berger et al. ("A Maximum Entropy Approach to Natural Language Processing," Computational Linguistics, 22 (1), pp. 39-68, 1996) teaches the use of context dependent word models.

Conclusion

Any response to this office action should be mailed to:

Commissioner of Patents and Trademarks
P.O. Box 1450
Alexandria, VA 22313-1450

or faxed to:

(703) 872-9314

Hand-delivered responses should be brought to:

Crystal Park II
2121 Crystal Drive
Arlington, VA.
Sixth Floor (Receptionist)

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dr. V. Paul Harper whose telephone number is (703) 305-4197. The examiner can normally be reached on Monday through Friday from 8:00 a.m. to 4:30 p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil, can be reached on (703) 305-9645. The fax phone number for the Technology Center 2600 is (703) 872-9314.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the Technology Center 2600 Customer Service office whose telephone number is (703) 306-0377.

V. Paul Harper

VPH/vph
March 30, 2004

RICHEMOND DORVIL
SUPERVISORY PATENT EXAMINER